NCAR

NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

Computational costs of WRF and WRF-Chem simulations

Negin Sobhani

National Center for atmospheric Research(NCAR)

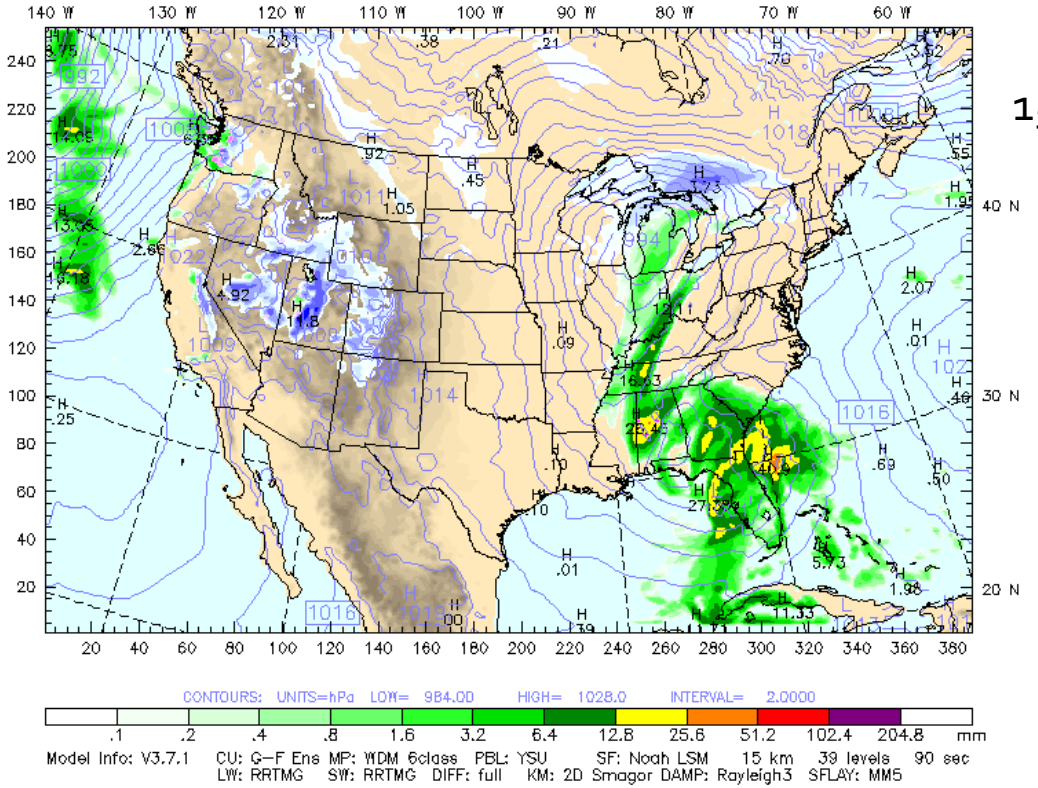
negins@ucar.edu

WRF is a numerical weather prediction system designed for both atmospheric research and operational forecasting.

Community model with **large user base**:

- More than 30,000 users
  in 150 countries

- WRF scalability and MPI parallelism
- Identifying hotspots and potential areas for improvement in WRF and WRF-Chem
- Profiling some chemistry options



15 km X15 km

~2 million gridpoints

Higher resolutions???

We NEED more cores (parallelism.)

General Considerations for Determining the Computational Costs

Question: "Is it possible to solve a problem with such-and-such resolution in a timely manner?"

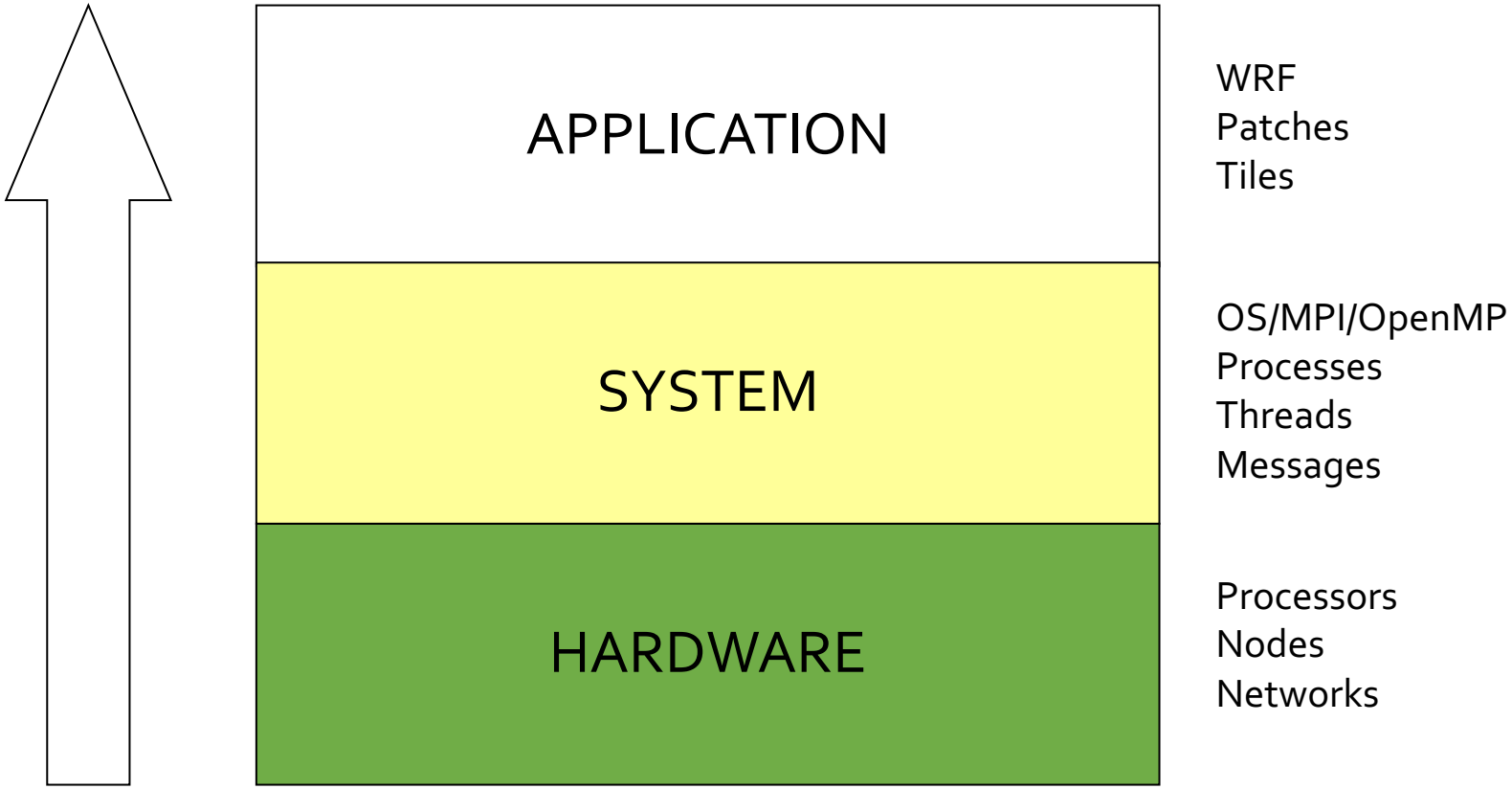
- What is my available computational power (allocation)?
- How much storage do I have?

- How quickly do I need a solution?(Operational vs. retro runs)
- What horizontal and vertical and vertical resolution for my purpose?
- What schemes? What is the purpose of my simulations (e.g. Dust or Biogenic)

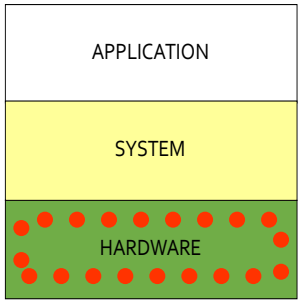
- How many cores?
  - If I use more cores I will have the results more quickly? (wrong)

- How large is any data set that you need to load?
- How much memory needs to be available for you to complete a run?

Computing Framework of WRF



From: Tutorial Notes: WRF Software Architecture by Michalakes and Gill



Parallel computing terms -- hardware

- **Processor**:
  - A device that **reads and executes instructions in sequence to produce perform operations on data** that it gets from a memory device producing results that are stored back onto the memory device
- **Node**: One memory device connected to one or more processors.
  - Multiple processors in a node are said to share-memory and this is "shared memory parallelism"
  - They can work together because they can see each other's memory
  - The latency and bandwidth to memory affect performance
- **Cluster**: Multiple nodes connected by a network
  - The processors attached to the memory in one node can not see the memory for processors on another node
  - For processors on different nodes to work together they must send messages between the nodes. This is "distributed memory parallelism"
- **Network:**
  - Devices and wires for sending messages between nodes
  - Bandwidth – a measure of the number of bytes that can be moved in a second
  - Latency – the amount of time it takes before the first byte of a message arrives at its destination

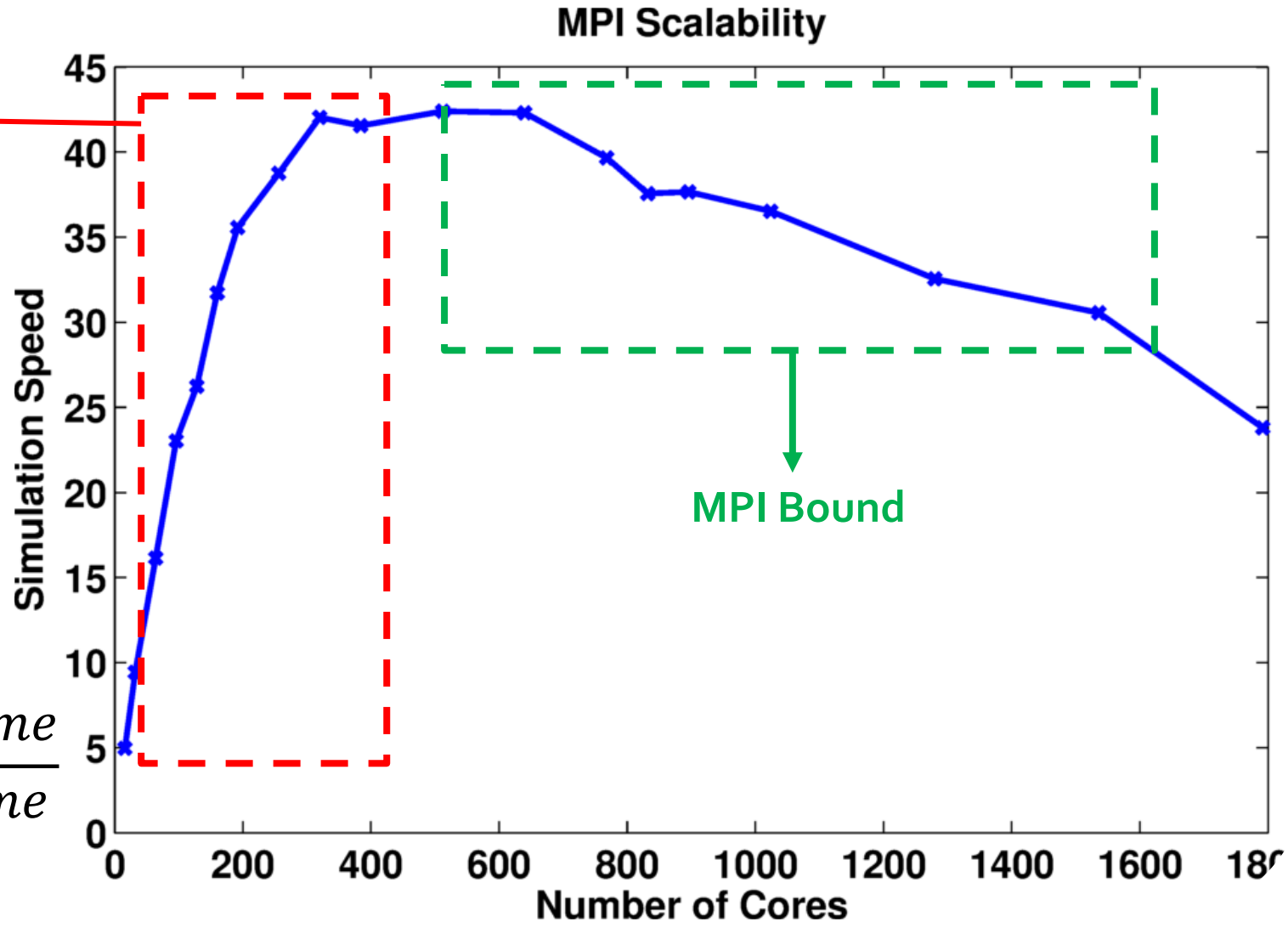
From: Tutorial Notes: WRF Software Architecture by Michalakes and Gill

Scalability Assessment (MPI Only)

500X500 horizontal grids

Simulation Speed is duration of simulation per wall clock time

$$Simulation\ Speed = \frac{Simulation\ Time}{WallClock\ Time}$$

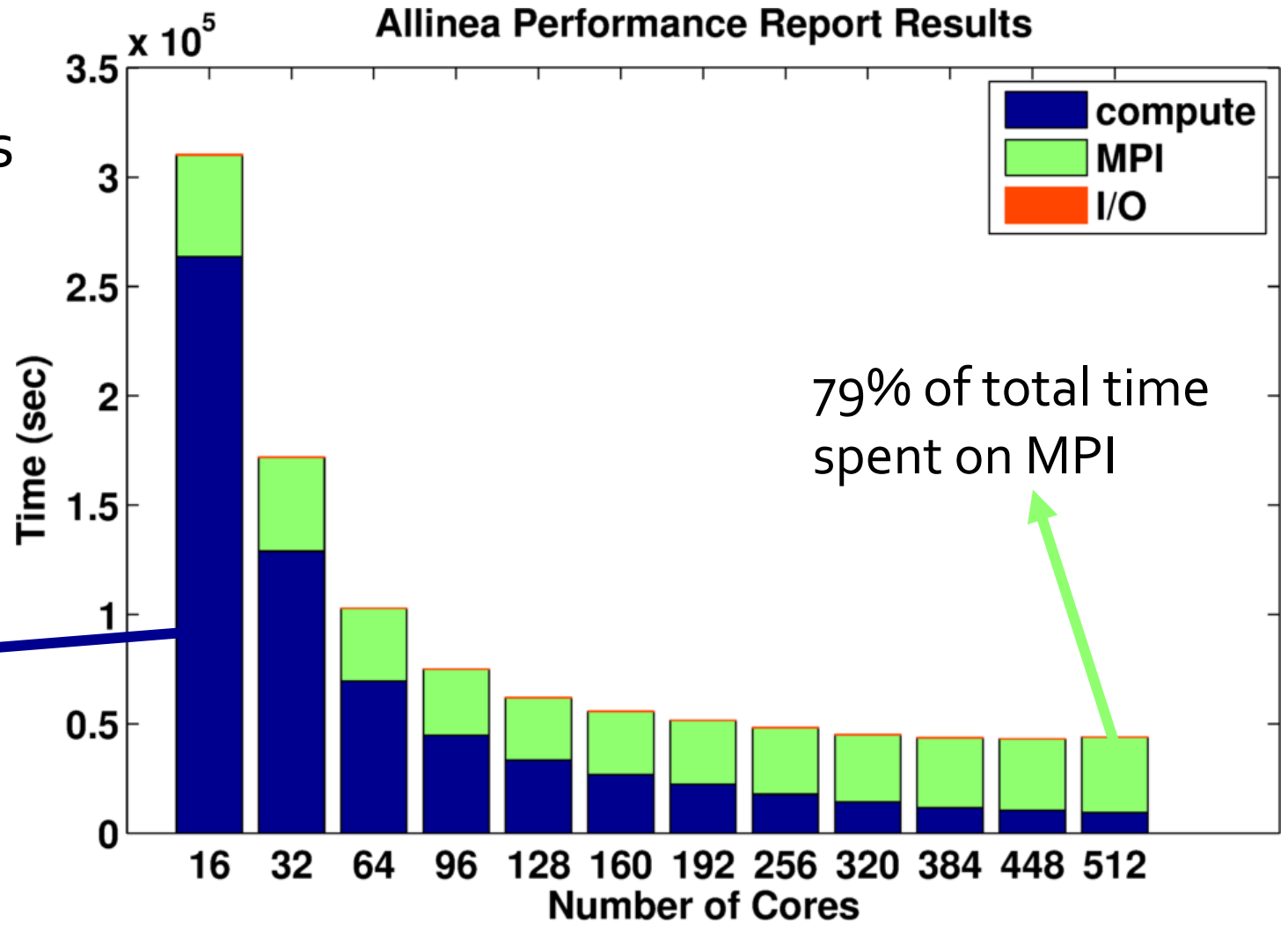


From: Sobhani et al. , AMS 2017

Scalability Assessment (MPI Only)

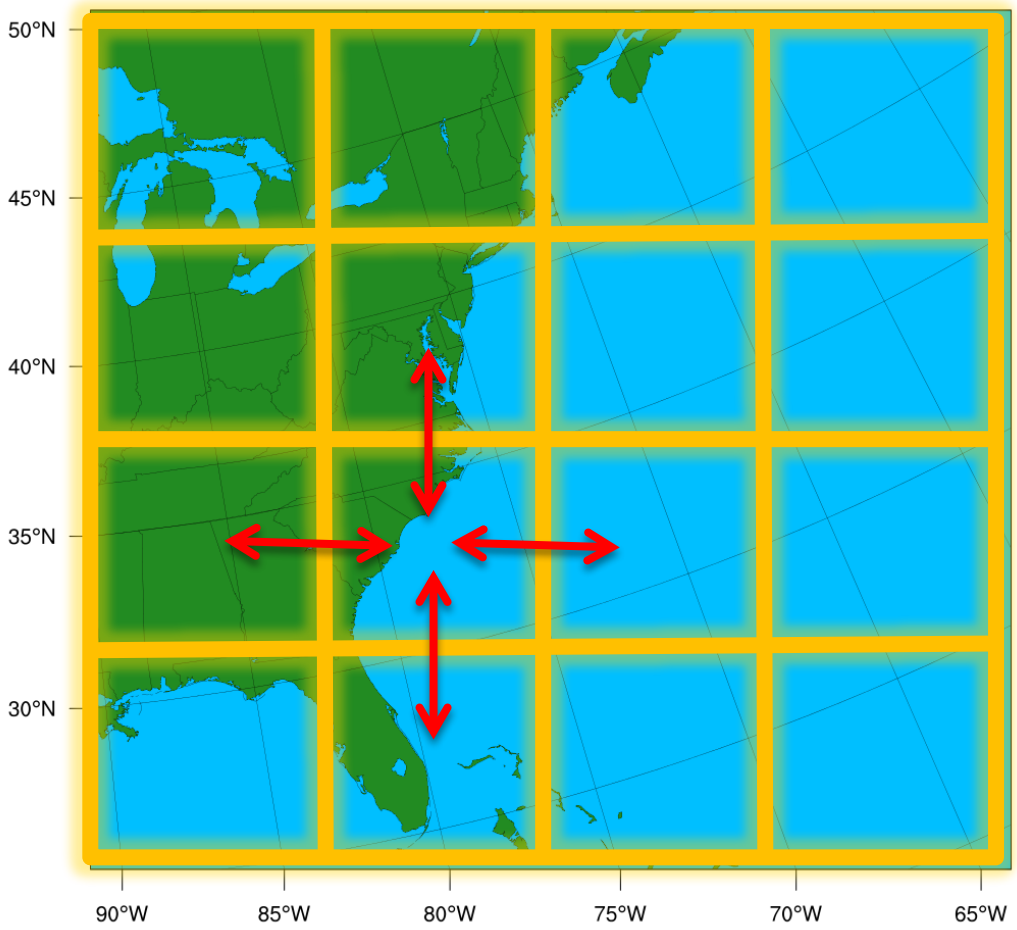
- Allinea Perfomance Reports
- Separate netcdf output file

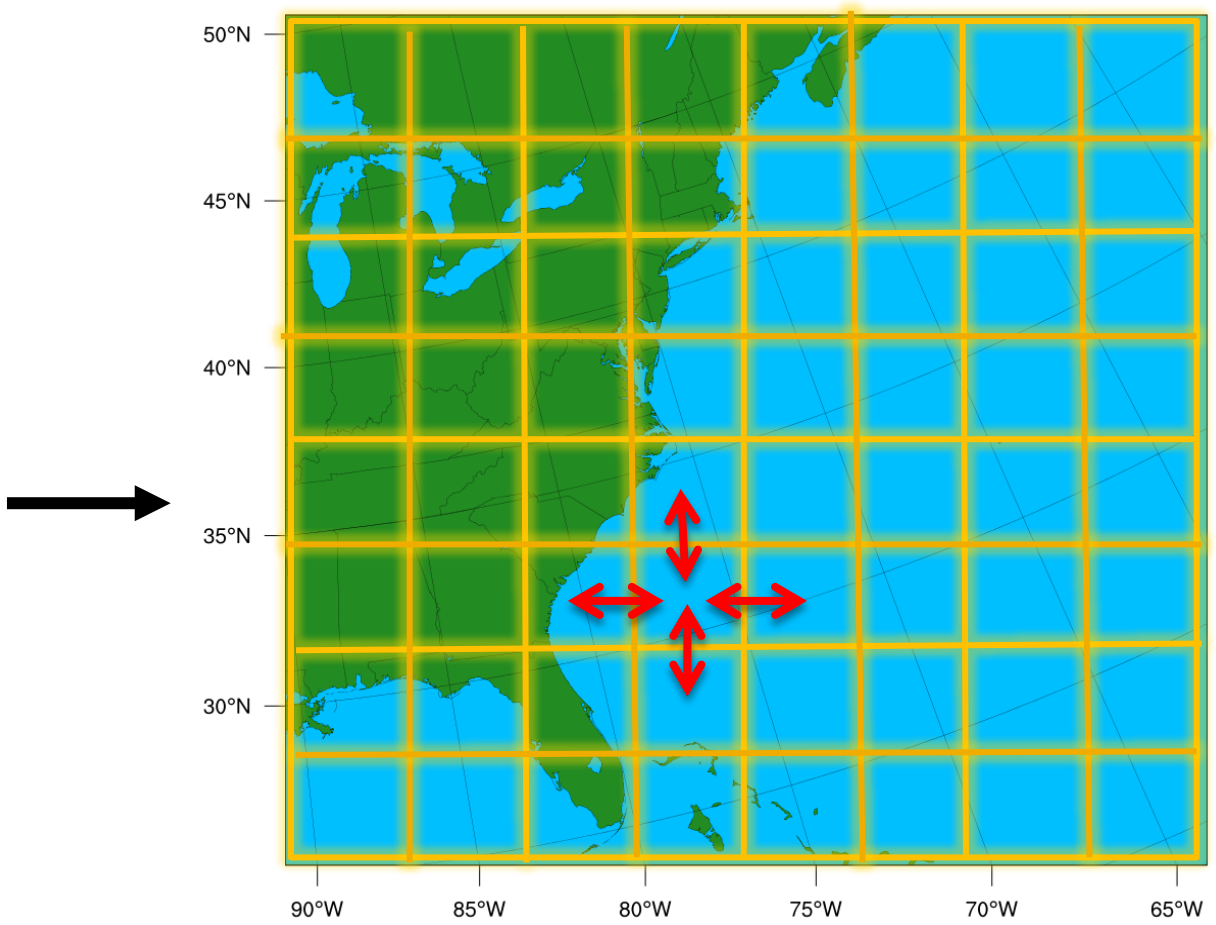
(io_form_history=102 in WRF namelist)



From: Sobhani et al. , AMS 2017

Domain Decomposition (MPI only)



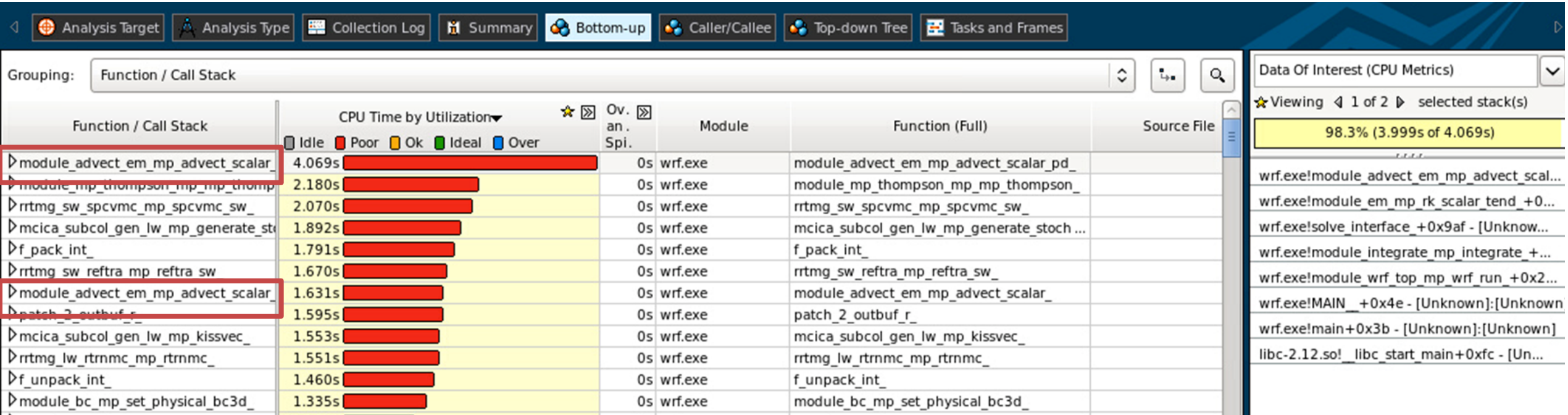


Per Grid :1/4 Computation and 1/2 MPI

From: Sobhani et al. , AMS 2017

Intel Vtune Amplifier XE

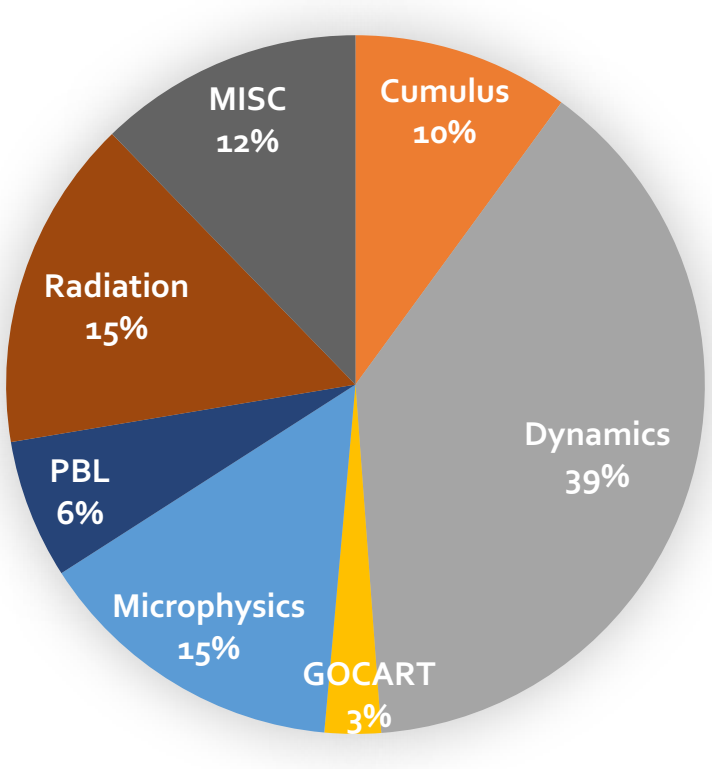
- Intel profiling and performance analysis tool
- Profiling includes stack sampling, thread profiling and hardware event sampling
- Collect performance statistics of different part of the code



What makes WRF-Chem expensive?

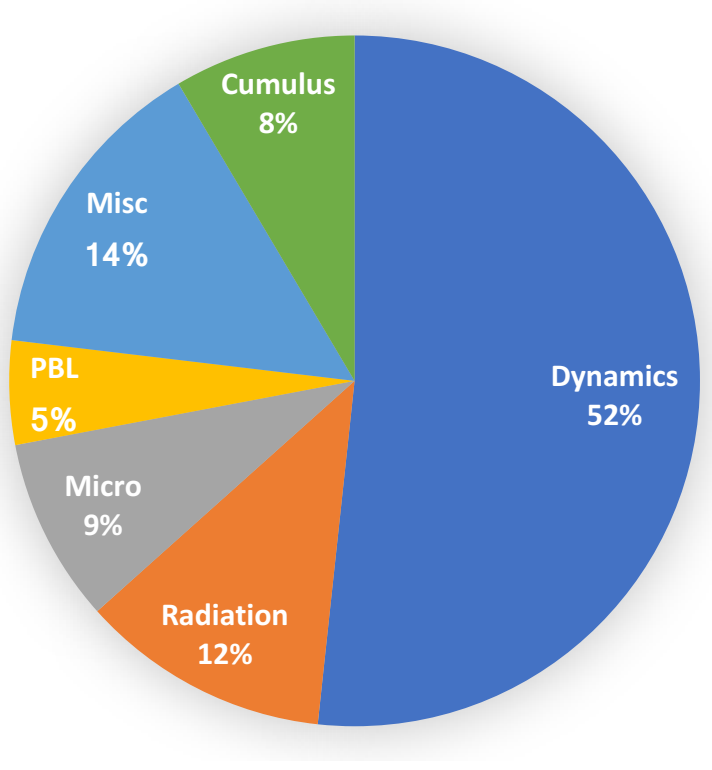
Question: How many species/tracers?

Example 01 – Dust
Chem_opt = 401
Dust Option = 1 GOCART



Total time: 132

Example 02 – Dust+Tracers
Chem_opt = 301
Dust Option: 01 GOCART

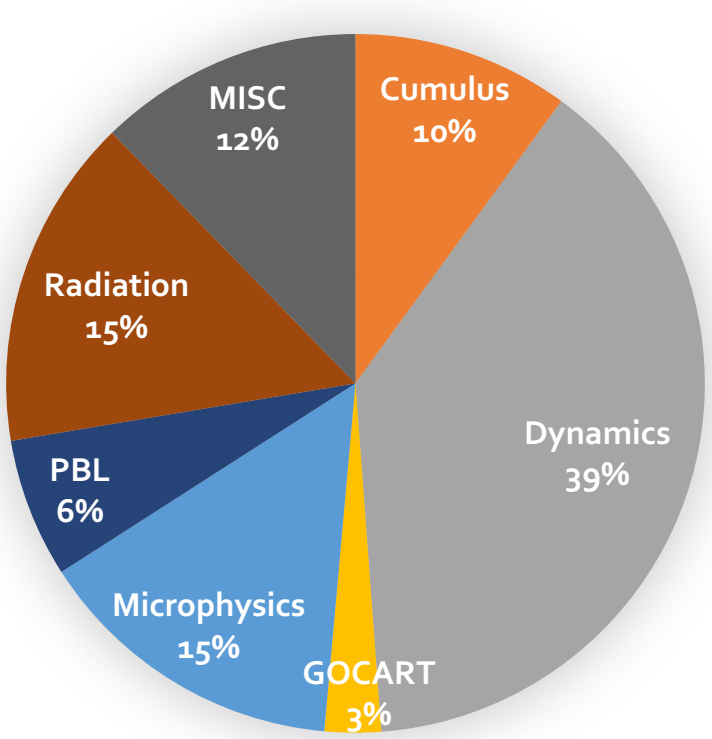


Total time: 203

What makes WRF-Chem expensive?

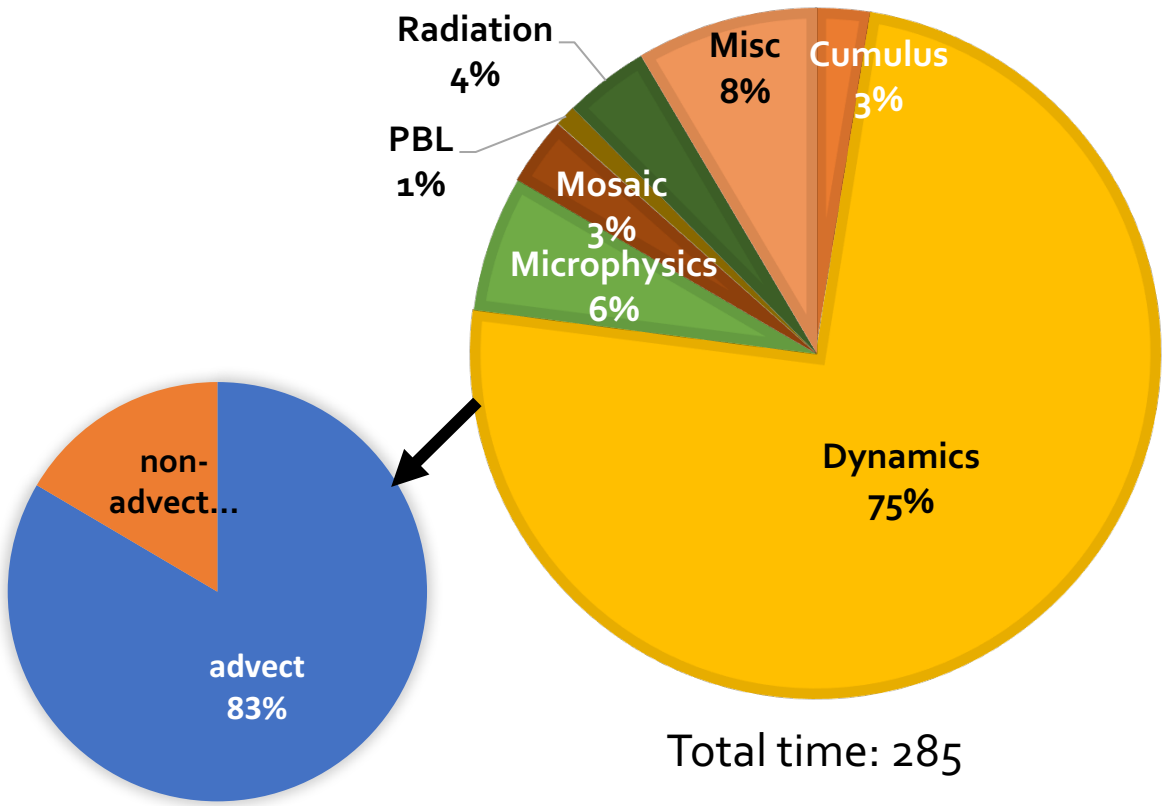
Question: How many species/tracers?

Example 01 – Dust
Chem_opt = 401
Dust Option = 1 GOCART



Total time: 132

Example 03 – Dust+ Tracers
Chem_opt = 8 -Tracer mode
MOSAIC 8 bin

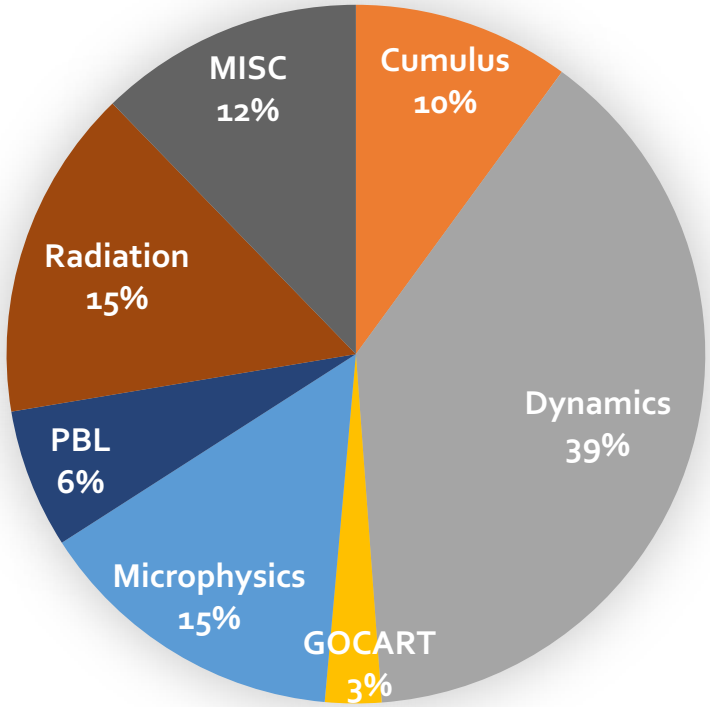


Total time: 285

What makes WRF-Chem expensive?

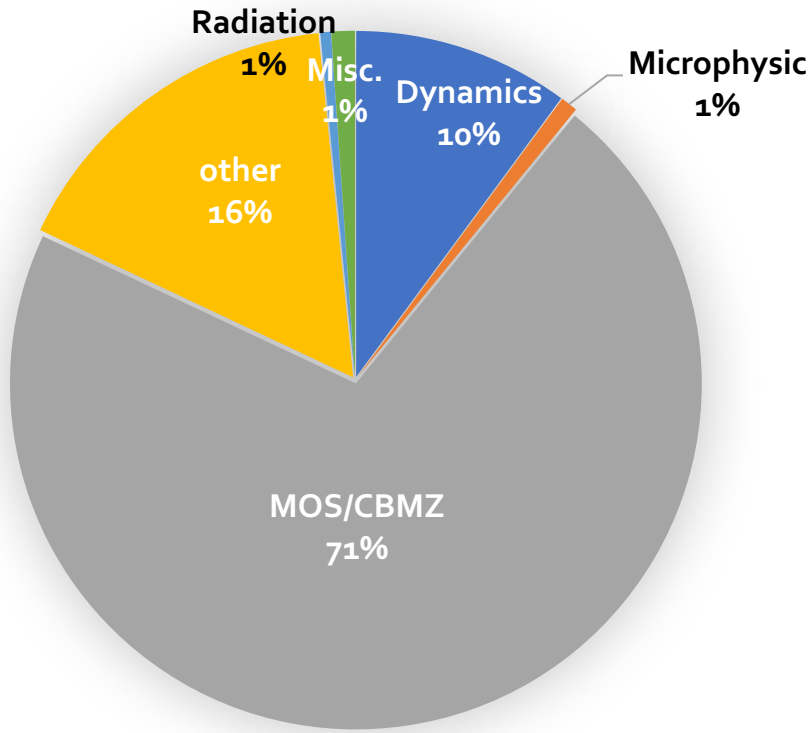
Question: Which Chemistry?

Example 01 – Dust
Chem_opt = 401
Dust Option = 1 GOCART



Total time: 132

Example 03 – Dust+ Tracers
Chem_opt = 8 - Chemistry
MOSAIC 8 bin



Total time: 2550

Transport in WRF and WRF-Chem

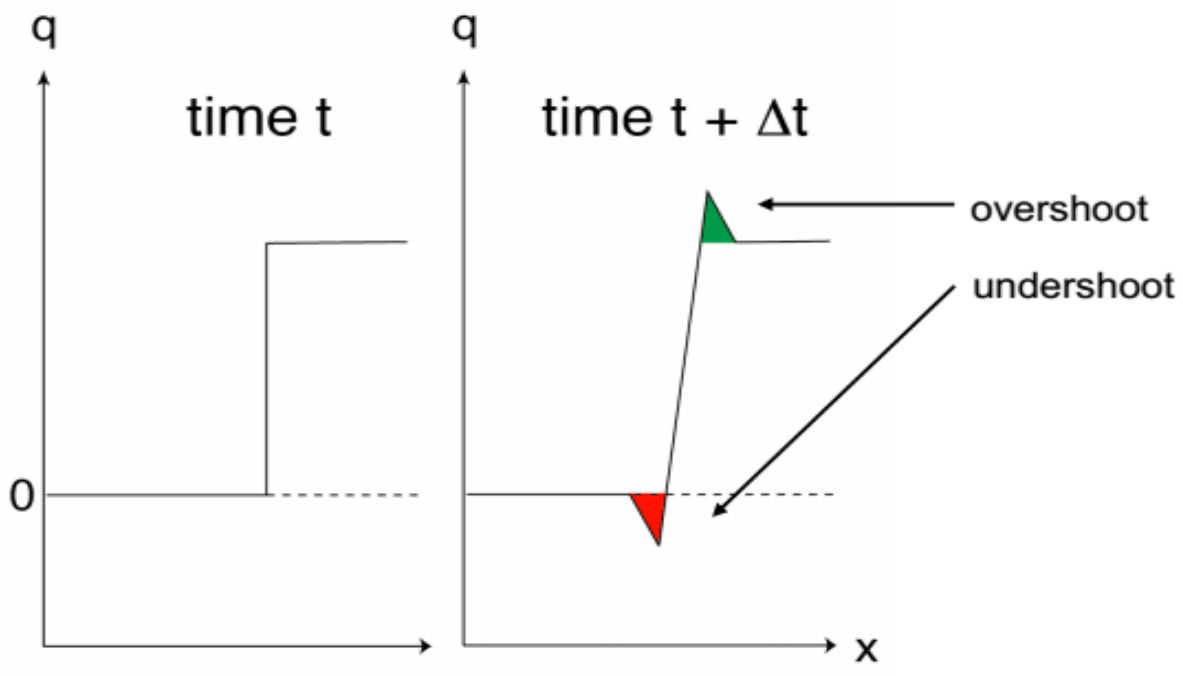
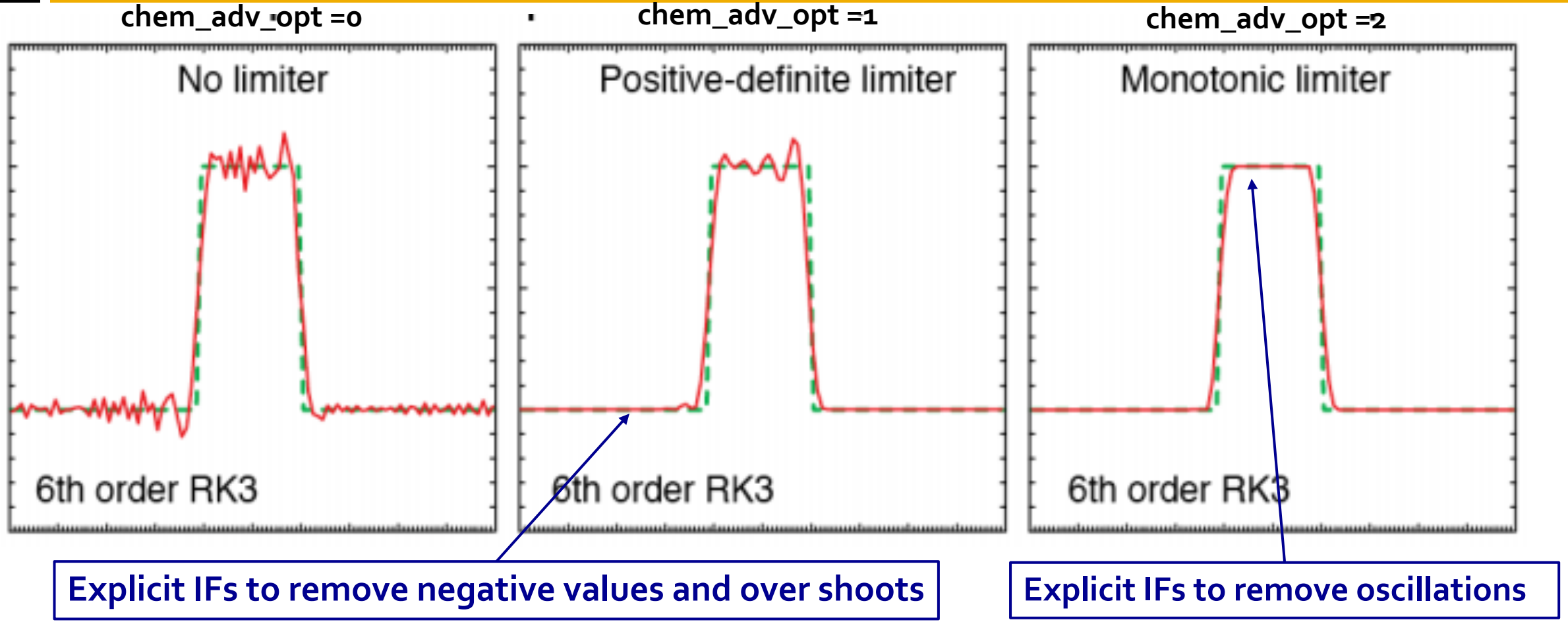


Figure from Skamarock and Dudhia 2012

- Until recently, many weather models did not conserve moisture because of the numerical challenges in advection schemes. → high bias in precipitation
- WRF-ARW is conservative but not all of the advection schemes are.
- This introduces new masses to the system.

Advection schemes can introduce both positive and negative errors particularly at sharp gradients.

Advection options in WRF

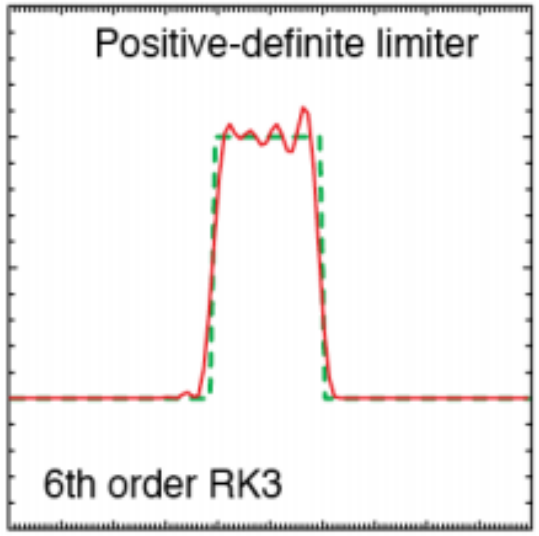


- **High number of explicit IFs are causing high branch mispredictions**

Figure from Skamarock and Dudhia 2012

Positive Define Advection Scheme Optimization

moist_adv_opt =1



Hotspot

Positive Definite Delimiter (32 lines)

- High Time
- High cache misses (both L1 and L2 Cache misses)
- High branch miss-prediction

Optimization Solution

Restructure and split the PD delimiter loop

- Increase vectorization
- Reduce cache misses

These optimizations are now included in the WRF original source code and available to public since V3.9 release.

Compiler	Optimization Flag	Loop Speed-up	Kernel Speed-up
Intel(v16.0.2)	-O3	100%	~17%
GNU (v6.1.0)	-Ofast	105%	~11%
PGI (v16.5)	-O3	35%	~4%

From Sobhani et al. , AMS 2017

Summary

- **Benchmarking:** Do some benchmark runs to estimate the number of core-hours you will need for each planned simulation.
- **Scaling:** Find the optimum number of nodes with scaling.
- **Profiling:** WRF with Intel Vtune XE, TAU tools, and Allinea MAP for identifying the hotspots of WRF
- **Optimizing:** the identified hotspots of different advection schemes for Intel, GNU , and PGI compilers
  - Significant speed-up of the advection schemes
- **Integration**: The changes to WRF advection schemes are approved by the WRF committee and **integrated** in the main WRF repository.